SYSTEMS AND METHODS FOR COMBINED BROWSING AND SEARCHING IN A DOCUMENT COLLECTION BASED ON INFORMATION SCENT

INCORPORATION BY REFERENCE

[0001] The following co-pending applications:

"SYSTEMS AND METHODS FOR IDENTIFYING USER TYPES USING MULTI-MODAL CLUSTERING AND INFORMATION SCENT", by E. Chi et al., Docket No. D/A0A28, files March 30, 2001 as U.S. Application Serial No. _____;

"SYSTEMS AND METHODS FOR PREDICTING USAGE OF A WEB SITE USING PROXIMAL CUES", by E. Chi et al., Attorney Docket No. D/A0A29, filed March 30, 2001 as U.S. Application Serial No. _____;

"SYSTEMS AND METHOD FOR INFORMATION BROWSING USING MULTI-MODAL FEATURES", by F. Chen et al., Attorney Docket No. D/99011, filed October 19, 1999, as U.S. Application Serial No. 09/421770;

"SYSTEM AND METHOD FOR PROVIDING RECOMMENDATIONS BASED ON MULTI-MODAL USER CLUSTERS", by H. Schuetze et al., Attorney Docket No. D/99197, filed October 19, 1999, as U.S. Application Serial No. 09/425038

"SYSTEM AND METHOD FOR QUANTITATIVELY REPRESENTING DATA OBJECTS IN VECTOR SPACE", by H. Schuetze et al., Attorney Docket No. D/99198, filed October 19, 1999, as U.S. Application Serial No. 09/421416;

"SYSTEM AND METHOD FOR IDENTIFYING SIMILARITIES AMONG DOCUMENTS IN A COLLECTION", by H. Schuetze et al., Attorney Docket No. D/99198Q1, filed October 19, 1999 as U.S. Application Serial No. 09/421767;

"SYSTEM AND METHOD FOR CLUSTERING DATA OBJECTS IN A COLLECTION", Schuetze et al., Attorney Docket No. D/99198Q2, filed October 19, 1999 as U.S. Application Serial No. 09/425039;

"SYSTEM AND METHOD FOR VISUALLY REPRESENTING THE CONTENTS OF A MULTIPLE DATA OBJECT CLUSTER", by H. Schuetze et al., Attorney Docket No. D/99198Q3, filed October 19, 1999, as U.S. Application Serial No. 09/421419;

"SYSTEM AND METHOD FOR INFERRING USER INFORMATION NEED IN A HYPERMEDIA LINKED DOCUMENT COLLECTION " by Ed Chi et al., Attorney

Docket No. D/99794, filed March 31, 2000, as U.S. Application Serial No. 09/540063; are each incorporated herein by reference in the entirety.

GOVERNMENT LICENSE PROVISION

[0002] The U.S. Government has a paid-up license in this invention and the right in limited circumstances to require the patent owner to license others on reasonable terms as provided for by the terms of Contract No. N00014-96-C-0097 awarded by the Office of Naval Research.

BACKGROUND OF THE INVENTION

1. Field of Invention

[0003] This invention relates to computer assisted search and retrieval systems and systems and methods for combined browsing and searching of a document collection or web site.

2. Description of Related Art

[0004] The ability to manage information is increasingly important in the modern information economy. As the reach of corporate information systems is extended to suppliers and customers, timely access to corporate information repositories becomes critical. Therefore, web site designers and information architects need to provide users with tools that facilitate efficient access to required information.

[0005] Users interact with conventional information systems to accomplish tasks using distinct modes of access. If a user is familiar with the terminology used in a particular domain, such as printers, the user is likely to know the keywords likely to retrieve relevant information required to complete the user's task. For example, the use of the word "multi-function" to describe machines that combine printer, scanner, copier and fax capabilities. The task of determining what "multi-function" printers exist may be accomplished using the search mode of an information system. The keywords are entered and the results are displayed as a list of documents containing the keywords. The user then selects the document that appears most relevant and reads each document presented.

[0006] However, in the search mode, if the user selects keywords that are too specific or not relevant to the subject matter, few if any documents will be selected and the user is given the false impression that no information exists. For

example, if "multi-function laser" were entered , the displayed documents would probably not include multi-function units that employed ink jet print output devices.

[0007] If the user selects keywords that are not specific enough, too many documents will be selected and the user will be overloaded with extraneous documents. Since few users will review the second or subsequent pages of a search request, the retrieval of large amounts of information tends to increase the user's cognitive overhead

[0008] If a user is unfamiliar with the subject area and therefore does not know the relevant keywords to generate a search query, a browsing mode of the information system is initiated. The user then identifies the relevant subject area specific keywords. Once the relevant keywords are identified, the user may enter the search mode and initiate a keyword search based on the information obtained from browsing.

[0009] The separation of the search and browse modes results in cognitive interruptions of the user's session. As relevant keywords are identified in the browsing mode, an interruption occurs as the user switches to the search mode to determine how well the keyword functions in narrowing the search mode results. A switch back to the browse mode may then occur. Using the browse/search process, a query is gradually developed that identifies the relevant information to accomplish the user's task. However constant switching between the search and browse modes consumes a great deal of the user's cognitive attention and requires considerable user training in developing search strategies.

[0010] In response some vendors of information systems have attempted to share previous user's browse paths under the assumption that many user's will have the same information requirements. Conventional systems such as IBM's SurfAid product and Alexa Internet's ToolBar 5.0 facilitate sharing of information obtained through a user's browsing mode experience. For example, Alexa Internet's Toolbar 5.0 system provides a customized toolbar that is added to the client browser. Using the Toolbar 5.0 product, Alexa Internet is able to compile information regarding a user's path in the browsing mode and makes suggestions of a next connection based on the similarity of the current path to accumulated historical browsing information. Similarly IBM's SurfAid product uses On-Line Analytical Processing methods to

provide a user with counts of other users following traversal paths in a browsing mode.

[0011] However, these conventional systems do not provide integration between the search mode and the browse modes. Also these conventional systems do not use information scent to determine relevancy of information tailored to the user using low cognitive overhead.

SUMMARY OF THE INVENTION

[0012] Therefore, the ability to determine high relevancy paths using information scent and to integrate the search and browse modes into a single interface would be useful.

[0013] The various exemplary embodiments of this invention provide systems and methods for combining browsing and searching of a document collection or web site using information scent.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows an exemplary embodiment of a system for combined browsing and searching in a document collection based on information scent according to this invention;

Fig. 2 show an expanded view of an exemplary embodiment of a system for combined browsing and searching in a document collection based on information scent according to this invention;

Fig. 3 is an exemplary flowchart of one embodiment of a method for combined browsing and searching in a document collection based on information scent according to this invention;

Fig. 4 shows a flowchart of an exemplary determination of information scent according to this invention;

Figs. 5 shows a first view of an exemplary document collection indicating the flows of information scent according to one embodiment of this invention;

Fig. 6 shows a second view of the exemplary document collection indicating the flows of information scent according to one embodiment of this invention;

Fig. 7 shows a third view of an exemplary document collection indicating the flows of information scent according to one embodiment of this invention;

Fig. 8 shows a first view of an exemplary hypermedia document according to one embodiment of the invention;

Fig. 9 shows a second view of an exemplary hypermedia document according to one embodiment of the invention;

Fig. 10 shows a third view of an exemplary hypermedia document according to one embodiment of the invention.

DETAILED DESCRIPTION OF EXEMPLARY EMBODIMENTS

[0014] Fig. 1 shows an exemplary embodiment of a system for combined browsing and searching based on information scent 100 according to this invention. Communications links 110 connect document server or web server 80 to the system for combined browsing and searching based on information scent 100.

[0015] In one exemplary embodiment according to this invention, access by the document or web browser of web-enabled computer 200, to the document server 80 over communications links 110 is mediated by the system for combined browsing and searching based on information scent 100. It will be apparent that the system can mediate access by acting as a proxy for access by web-enabled computer 200 to the web server 80 or in various other embodiments, by incorporating the system for combined browsing and searching based on information scent 100 within a web-enabled computer as shown in 300.

[0016] The document or web server 80 makes available documents or web pages to client document or web browsers over communications links 110. When the web browser of web-enabled computer 200 starts, the user is presented with a display that allows keyword information and an initial document or starting location in the document collection to be entered. For example, the keyword might be "remote diagnostics" and the initial document or starting location might be the ABC Corp web site at "www.abc-corp.com".

[0017] This information is then transferred to the system for combined browsing and searching based on information scent 100. The system for combined browsing and searching based on information scent 100 contains previously obtained topology and content information for each site for which access is mediated. The system for combined browsing and searching based on information scent 100 determines the document requested by the web browser of web-enabled computer 200 and initiates a request for the document or web page from document server or web server 80 over communications links 110. The retrieved document or web page is then analyzed based on the user's information goal indicated by the user-entered

keywords to determine which of the connections or links on the retrieved document or web page provide the greatest flow of information scent for documents or web pages relevant to the keywords. Display attributes are then synthesized for the connections indicating the greatest flow of information scent. The display attributes can be any sensible characteristic. However, in one of the various exemplary embodiments, the font size characteristic of the connections or links of retrieved documents or web pages are synthesized to indicate increasing relevancy by increasing the font size. It will be apparent that other types of synthesized display characteristics may be used. For example, changing font color from red to green, increasing bolding of text, making the font of more relevant connections or links italics or any other known or later developed method of indicating flow of information scent may be used in the practice of this invention.

[0018] The retrieved document or web page with connections or links having the synthesized display attribute is then transferred from the system for combined browsing and searching based on information scent 100 to the web browser of web-enabled computer 200.

[0019] The web-enabled computer 300 includes a modified document or web browser that integrates the functionality of the system for combined browsing and searching based on information scent 100 into web-enabled computer 300. It will be apparent that the system for combined browsing and searching based on information scent 100 may be implemented as a server mediating access for multiple computers, as a routine or software manager on computer 300 or any other combination without departing in the practice of this invention.

[0020] Fig. 2 shows an exemplary embodiment of a system for combined browsing and searching based on information scent 100. The system for combined browsing and searching based on information scent 100 comprises a controller circuit 10; a memory circuit 14; a topology determining circuit 16; a content determining circuit 18; a user keyword determining circuit 20; an information scent determining circuit 22; a document distance determining circuit 24; a browser request determining circuit 26; a browser request retrieving circuit 28; a display attribute synthesizing circuit 30; a relevant document determining circuit 32; a relevant document path determining circuit 34; an input/output circuit 12 connected through communications link 110 to document or web server 80 providing access to document collection or

web site 90. The communications link 110 also provides a connection to input/output circuit 212 of web-enabled computer 200. Web enabled computer 200 comprises a controller circuit 210; an input/output circuit 212 for connecting to communications links 110; a memory circuit 214 and browser circuit 216.

[0021] The controller circuit 10 activates topology determining circuit 16 to retrieve topology information about web site 90 and store the information in memory circuit 14. The topology may be determined by traversing the site and identifying connections or links between documents or web pages. For example, starting at a first document or web page, the documents or web pages connected to, or linked to, the first document or web page are determined. Information indicating an association between the first document or web page and the reachable documents or web pages is stored in a topology data structure. It will be apparent that a topology data structure may include a topology matrix, a topology adjacency list or any other known or later developed technique of storing topology information about the documents or web pages in the document collection or web site.

[0022] The controller circuit 10 activates the content determining circuit 18 to retrieve content information concerning each document or web page in the document collection or web site 80 and store the content information in memory circuit 14. It will be apparent that the content information may be obtained at the same time as the topology of the document collection or web site 80 is determined by topology determining circuit 16 or may be determined after the topology has been determined. The content information may be determined using any known or later developed technique of content determination such as web crawling.

[0023] The content of each of the documents or web pages making up the document collection are determined. The words on each document or web page are added to a word / document frequency matrix. The weights of the words are determined and a weighted word document frequency matrix is created. The weighting may use term frequency/inverse document frequency, log of the term frequency, 1+ ($\log_{10}$ of the term frequency) or any other known or later developed technique of weighting.

[0024] The controller circuit 210 of web-enabled computer 200 activates the browser circuit 216 which generates a request for an initial document or home page. In response to the initial document or home page request from browser circuit 216 of

web enabled computer 200, the system for combined browsing and searching based on information scent 100 activates the user keyword determining circuit 20 to determine the keywords describing the user's information goal or task. The keywords can be used to describe the overall task in general terms to facilitate browsing of the document collection or web site 80 or may be highly specific words in order to more quickly determine the relevant links. The greater the number of words specified, the more the results will reflect a search mode retrieval from the document collection. Similarly the fewer the number of keywords entered, the more the results will reflect a browsing mode retrieval from the document collection or web site.

[0025] When an initial document or web page is requested by the browser circuit 216 of web-enabled computer 200, the controller circuit 210 sends the browser circuit document or web page request through input/output circuit 212 over communications link 110 through the input/output circuit of the system for combined browsing and searching based on information scent 100 to the browser request determining circuit 26. The controller circuit 10 of the system for combined browsing and searching based on information scent 100 activates the browser request retrieving circuit 28 which retrieves the requested document from document or web server 80 and stores the requested document or web page in memory circuit 14.

[0026] The controller circuit 10 of the system for combined browsing and searching based on information scent 100 activates the relevant document determining circuit 32. The relevant document determining circuit 32 determines the relevant documents in the document collection or web site using a search. The search may be any known or later developed search capable of indicating documents base don the user keywords. The controller circuit 10 of the system for combined browsing and searching based on information scent 100 activates the relevant document path determining circuit 34 to determine the path between the current document or web page and each relevant document or web page. The controller 10 then transfers the relevant document path to the information scent determining circuit 22 to determine an information scent associated with the relevant document path . The information scent determining circuit 22 determines the information scent for each connection or link on the retrieved page based on the user keywords and the previously retrieved topology and content of the document collection or web site. The determination of information scent for a document collection or web site may use the techniques

described in "SYSTEM AND METHOD FOR INFERRING USER INFORMATION NEED IN A HYPERMEDIA LINKED DOCUMENT COLLECTION " by Ed Chi et al., Attorney Docket No. D/99794, filed March 31, 2000, as U.S. Application Serial No. 09/540063; incorporated herein by reference in its entirety, or any other known or later developed technique for determining information scent in a document collection of web site. Information scent according to the system for combined browsing and searching based on information scent 100 flows in the direction opposite to that of the connections or links in document collection or web site. That is, the information flows backward over a link to provide cue information as to what can be found at the connected to or linked to end of the link.

[0027] The controller 10 of the system for combined browsing and searching based on information scent 100 then activates the document distance determining circuit 24 to determine how far the relevant documents are from the current document or web page in the web site topology. The current document or web page is the current document or web page currently being retrieved. The distance between pages may be calculated by analyzing the topology information of the document collection or web site stored in memory circuit 14 or any other technique of determining the number of documents or pages between the relevant documents and the current document . The topology information may be stored in a matrix, an adjacency list or any other known or later developed structure for storing the relationship between documents or web pages.

[0028] The controller circuit 10 then adjusts the determined information scent for each current connection based on the determined document distance. The controller circuit 10 adds the determined information scent for connections leading to relevant documents or pages thereby yielding more information scent or larger scent conduits.

[0029] The controller circuit then activates the display attribute synthesizer to re-write the stored document or web page by synthesizing a display attribute based on the determined scent information. For example, a display attribute such as a font size of the connection or link in the retrieved document stored in memory circuit 14 may be synthesized. In various exemplary embodiments according to this invention, the display attribute may change to indicate the amount of scent information associated with a connection or link. Any type of visual, auditory, tactile, olfactory or taste

display attribute known or later developed may be used in the practice of this invention. In various other exemplary embodiments of this invention, graphic images may also be used as connections or links. Display attributes for graphic image connections or links may include but are not limited to adding border color around the image, adding a hue saturation to the image or any other known or later developed technique of indicating changes.

[0030] The re-written document or web page is then transferred through input/output circuit 12 over communications link to the input/output circuit 212 of web-enabled computer 200. The browser circuit 216 of web-enabled computer 216 is then activated to display the re-written document or web page with synthesized display attributes indicating the information scent for each connection or link.

[0031] Fig. 3 is a flowchart of an exemplary embodiment of a method for combined browsing and searching based on information scent 100 according to this invention. The process starts at step S10 and immediately continues to step S20. In step S20 the topology and content of the document collection or web site is determined. Control then continues to step S30.

[0032] In step S30 the user keywords are determined. The user keywords may determined by prompting the user for the keywords using a pop-up dialog box, entry via a text field, voice input, or already stored user profiles, or any other known or later developed techniques. Control then continues to step S40 where the document requested by a user's browser is determined.

[0033] Control then continues to step S50 where the requested document is retrieved from the document collection and stored. In step S55, a search is performed in the document collection based on the user keywords. The relevant documents most closely matching the keywords are then identified and relevant document paths determined to each document. In step S60, the information scent associated with each of the determined relevant document paths is determined and an information scent vector is returned. The information scent vector indicates the relevancy of the associated connection or link for retrieval of the relevant document. The information scent is then adjusted based on a determination of how far the relevant document or web page is from the current document or page. The information scent vectors associated with a connection or link may be totaled to indicate the relevant strength of information scent associated with the connection or link. Information scent is added

together for connections or links leading to multiple paths that in turn lead to multiple relevant pages. Thus an adjustment of the information scent based on the topology of the document collection or web site is determined. Control then continues to step S70.

[0034] In step S70 display attributes are synthesized based on the determined information scent. For example, a font size or color may change based on the how well the information scent for a connection or link relates to the user keywords. The re-written document or web page containing the synthesized display attributes is then sent to the browser and control continues to step S80.

[0035] In step S80 a determination is made as to whether the user has requested that the process end. For example, the user may request the termination of the process by pressing the ESC key on the keyboard or by selecting a menu item or any other known or later developed technique of indicating that termination is desired.

[0036] If the determination step S80 determines that the user wishes to end the process, control continues to step S120 and the process ends. Otherwise control continues to step S90 where a determination is made whether a new document has been requested by the browser..

[0037] If a new document has been requested, control continues to step S40 where the document requested is determined. If a new document has not been requested, control continues to step S100.

[0038] In step S100, a determination is made whether the user has entered new user keywords. The user may enter new user keywords to increase the amount of information used in determining the information scent. If the determination in step S100 determines that no further user keywords are to be entered, control continues to step S80 and the process continues. Otherwise, if it is determined in step S100 that new keywords are to be entered, control continues to step S110.

[0039] In step S110 new user keywords are determined. For example, the user may enter keywords in a dialog box, enter text in a text entry field, select from a drop down list or any other known or later developed technique for entry or determination of user keywords. Control then continues to step S70 and the process repeats until the determination is made in step S80 that the user has requested that the process be ended at which point control continues to step S120 and the process ends.

[0040] Fig. 4 shows a flowchart of an exemplary method of determining information scent based on a relevant document path according to this invention. The process starts at step S400 and continues to step S410.

[0041] In step S410, the first relevant path to a relevant document is selected. In the exemplary embodiment, the relevant document path is determined using any known or later developed type of search to identify relevant documents based on the user keywords as described above with respect to step S55 of Fig. 3. Control then continues to step S420 where the content information for the document collection or web site is determined.

[0042] In the exemplary embodiment according to this invention, the content information is obtained from the stored content information determined in step S20 of Fig. 3. However, it will be apparent that any method of obtaining the content information may be used such as providing the content information as a parameter to the process of inferring user information need or by re-determining the content information as required. Control then continues to step S430 where the topology of the document collection or web site, is determined.

[0043] As discussed above it will be apparent that any method of obtaining the topology information may be used such as providing the topology information as a parameter to the process of inferring user information need, re-determining the topology information as required and/or retrieving the topology information stored in memory by step S20 of Fig. 3. Control then continues to step S440 where the document path position weighting and document access weighting are determined for the documents in the selected relevant document path. Control then continues to step S450.

[0044] In step S450, a weighted content data store is determined. The weighted content data structure may be a word x document matrix, a word x document adjacency list or any other known or later developed technique for storing the content information about the document collection or web site page. Control then continues to step S460.

[0045] In step S460 spreading activation according to the following formulas (1-2) is applied to generate initial document vector A.

$$A(1) = \text{ALPHA} * \text{Topology Matrix} * E \qquad (1)$$

$$A(t) = \text{ALPHA} * \text{Topology Matrix} * A(t-1) + E \qquad (2)$$

The formula is applied t number of times where the matrix W reflects the weighted content matrix and vector E reflects the relevant document path. The value ALPHA reflects the probability a user will click through to a document or web page and therefore ranges between 0 and 1. Control then may continue to optional step S470, or may in alternative embodiments according to this invention, continue to step S480.

[0046] In optional step S470, the document vector A is multiplied by the weighted content matrix to determine the information scent based on the relevant document path to create an information scent keyword vector. The most relevant keyword information is then indicated by higher number entries in the information keyword vector position. Control then continues to step S480 where the process ends and control is returned to the calling step S60 of Fig. 3.

[0047] Fig. 5 shows a first view of an exemplary document collection indicating the flows of information scent according to one embodiment of this invention. The initial or starting document or web page 91 is shown with the relevant user keywords "remote diagnostics". The flow of the information scent is represented by the width of the lines connecting each document or page and a number. Connections or links are shown underlined and documents or pages that have been determined relevant are shown hatched. The user keywords "remote diagnostics" render the documents or web pages "features remote diagnostics" 95 and "remote diagnostics" 96 the most relevant. Since the flow of the information scent occurs in the opposite direction to the links, and scent is additive, the "features remote diagnostics" 95 and "remote diagnostics" 96 documents add to provide a scent value of 4 from "features remote diagnostics" 95 to "AB 4411 copier features" 94. However, the scent also diminishes with distance, therefore, the scent between "AB 4411 copier features" 94 and "AB 4411/AB 5001" 93 has a value of 3. The scent diminishes again between "AB 4411/AB 5001" 93 and "digital copiers/color copiers/back" 92. Finally between "digital copiers/color copiers/back" 92 and "copiers/fax machines/other" 91 the value is only 1.

[0048] On the other hand, the scent value between "remote diagnostics" 96 and "maintenance" 97 is 3, and the scent between "maintenance and "copiers/fax machines/other" 91 is 2. Therefore a synthesized display attribute for the "other" connection or link will be made.

[0049] Fig. 8 shows an exemplary hypermedia document according to one embodiment of this invention which indicates an exemplary synthesized display attribute.

[0050] Fig. 6 shows a second view of an exemplary document collection indicating the flows of information scent according to one embodiment of this invention. The initial or starting document or web page 91 is shown with the relevant user keywords "AB 5001" reflecting the name of a copier. The user keywords "AB 5001" render the documents or web pages "AB 4411/AB5001" 93 the most relevant document. The information scent flows back towards "digital copiers/color copiers/back" 92 with a value of a 4 and from there to "copiers"/fax/machines/other" with a value of 3.

[0051] On the other hand, the scent value between "remote diagnostics" 96 and "maintenance" 97 is 1, and the scent between "maintenance and "copiers/fax machines/other" 91 is 1. Therefore a synthesized display attribute for the "copiers" connection or link will be made based on the determined information scent of 3.

[0052] Fig. 9 shows a second view of an exemplary hypermedia document according to one embodiment of this invention. The second view shows an exemplary synthesized display attribute for "copier products".

[0053] Fig. 7 shows a third view of an exemplary document collection indicating the flows of information scent according to one embodiment of this invention. The initial or starting document or web page 91 is shown with the relevant user keywords "AB 4411 copier features". The user keywords "AB 4411 copier features" render the documents or web pages "AB 4411/AB 5001" 93 and "AB 4411 copier features" 94 the most relevant. Since the flow of the information scent occurs in the opposite direction to that of the links, and information scent is additive, the "AB 4411 copier features" 94 and "AB 4411/AB 5001" 93 documents add to provide a scent value of 5 between "AB 4411/AB 5001" 93 and "digital copiers/color copiers/back" 92. However, the scent also diminishes with distance, therefore, the scent between "digital copiers/color copiers/back" 92 and "copiers/fax/machines/other" 91 has an attenuated value of 4 due to the distance from the two relevant documents or web pages.

[0054] On the other hand, the scent value between "remote diagnostics" 96 and "maintenance" 97 is 1, and the scent between "maintenance and "copiers/fax

machines/other" 91 is 1. Therefore a synthesized display attribute for the "copiers" connection or link will be made based on the determined information scent of 4.

[0055] Fig. 10 shows a third view of an exemplary hypermedia document according to one embodiment of this invention. The third view shows an exemplary synthesized display attribute for "copier products".

[0056] In the various exemplary embodiments outlined above, the system for combined browsing and searching based on information scent 100 can be implemented using a programmed general purpose computer. However, the system for combined browsing and searching based on information scent 100 can also be implemented using a special purpose computer, a programmed microprocessor or microcontroller and peripheral integrated circuit elements, an ASIC or other integrated circuit, a digital signal processor, a hardwired electronic or logic circuit such as a discrete element circuit, a programmable logic device such as a PLD, PLA, FPGA or PAL, or the like. In general, any device, capable of implementing a finite state machine that is in turn capable of implementing the flowcharts shown in Figs. 3-4 can be used to implement the system for combined browsing and searching based on information scent 100.

[0057] Each of the circuits 10-34 of the system for combined browsing and searching based on information scent 100 outlined above can be implemented as portions of a suitably programmed general purpose computer. Alternatively, circuits 10-34 of the system for combined browsing and searching based on information scent 100 outlined above can be implemented as physically distinct hardware circuits within an ASIC, or using a FPGA, a PDL, a PLA or a PAL, or using discrete logic elements or discrete circuit elements. The particular form each of the circuits 10-34 of the system for combined browsing and searching based on information scent 100 outlined above will take is a design choice and will be obvious and predicable to those skilled in the art.

[0058] Moreover, the system for combined browsing and searching based on information scent 100 and/or each of the various circuits discussed above can each be implemented as software routines, managers or objects executing on a programmed general purpose computer, a special purpose computer, a microprocessor or the like. In this case, the system for combined browsing and searching based on information scent 100 and/or each of the various circuits discussed above can each be implemented as one or more routines embedded in the communications network, as a

resource residing on a server, or the like. The system for combined browsing and searching based on information scent 100 and the various circuits discussed above can also be implemented by physically incorporating the system for combined browsing and searching based on information scent 100 into a software and/or hardware system, such as the hardware and software systems of a document server, web server or electronic library server.

[0059] As shown in Fig. 3, the memory circuits 14 and 214, can be implemented using any appropriate combination of alterable, volatile or non-volatile memory or non-alterable, or fixed, memory. The alterable memory, whether volatile or non-volatile, can be implemented using any one or more of static or dynamic RAM, a floppy disk and disk drive, a write-able or rewrite-able optical disk and disk drive, a hard drive, flash memory or the like. Similarly, the non-alterable or fixed memory can be implemented using any one or more of ROM, PROM, EPROM, EEPROM, an optical ROM disk, such as a CD-ROM or DVD-ROM disk, and disk drive or the like.

[0060] The communication links 110 shown in Figs. 1-2 can each be any known or later developed device or system for connecting a communication device to the system for combined browsing and searching based on information scent 100, including a direct cable connection, a connection over a wide area network or a local area network, a connection over an intranet, a connection over the Internet, or a connection over any other distributed processing network or system. In general, the communication link 110 can be any known or later developed connection system or structure usable to connect devices and facilitate communication

[0061] Further, it should be appreciated that the communication link 110 can be a wired or wireless link to a network. The network can be a local area network, a wide area network, an intranet, the Internet, or any other distributed processing and storage network.

[0062] While this invention has been described in conjunction with the exemplary embodiments outlines above, it is evident that many alternatives , modifications and variations will be apparent to those skilled in the art. Accordingly, the exemplary embodiments of the invention, as set forth above, are intended to be illustrative, not limiting. Various changes may be made without departing from the spirit and scope of the invention.